Advanced Topics in Robust Control

Volume II: Invariants and Structured Singular Values

Final Technical Report
Period Covered: April 1982 to July 1990
Contract No. N00014-82-C-0157

August 1, 1990

Prepared by

Honeywell

Systems and Research Center
3660 Technology Drive
Minneapolis, Minnesota 55418

Prepared for

Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217-5000

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
August 1, 1990

3. REPORT TYPE AND DATES COVERED
Final Technical Report , Vol. 2 4/82 to 8/90

4. TITLE AND SUBTITLE
Advanced Topics in Robust Control
Volume 2: Invariants and Structured Singular Values

5. FUNDING NUMBERS
Contract No. N00014-82-C-0157

6. AUTHOR(S)
Blaise Morton, John Doyle, Andrew Packard, and Allen Tannenbaum

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Honeywell Systems and Research Center
3660 Technology Drive
Minneapolis, Minnesota 55418

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N. Quincy Street
Arlington, Virginia 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The topic is the theory of structured singular values in robust control analysis. Classical methods of invariant theory are applied to the problem of computing the structured singular value for the nonrepeated diagonal complex block structure. An algorithm for computing the structured singular value for the four-block problem is described.

14. SUBJECT TERMS
Control Systems ~~Analysis~~, Robust Control, Multivariable Control, Structured Singular Values, Invariant Theory

15. NUMBER OF PAGES
38

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

TABLE OF CONTENTS

INVARIANTS AND STRUCTURED SINGULAR VALUES

Introduction

The theme of the ONR-sponsored research program at Honeywell for the last eight years was the mathematical theory of robust control. The theoretical novelty that fueled the technical effort was the structured singular value (SSV). The result was a new theory of robust control based on the structured singular value. Control analysis and design tools based on this theory have gained widespread industry acceptance.

Quoting Gary Balas of MUSYN:

> To gauge the control community's interest in these topics, a short course offered in September 1989 by MUSYN Inc. titled, "Theory and Applications of Robust Multivariable Control," was attended by over 45 people from industry, academia and government laboratories. The course was offered again in March, 1990 at NASA Langley Research Center to 25 NASA and Air Force engineers. These researchers are interested in applying these techniques to topics such as: flight control, flutter suppression and vibration attenuation of flexible structures. In the summer of 1990 the course will be offered in Cambridge, England Delph, Netherlands and Pasadena, California. It is believed that over 200 people will attend the robust multivariable short course in 1990.
>
> Currently, H_∞ control design techniques and μ-analysis and synthesis methods are being used to design flight control systems, vibration attenuation control laws for flexible structures, and missile autopilots. Johns Hopkins University Applied Physics Laboratory (JHUAPL), China Lake Naval Weapons Center and Dahlgren Naval Weapons Center are applying such methods to design of robust control laws for missile autopilots. JHUAPL, in addition, is investigating the application of the the μ-synthesis methodology to the design of gain-scheduled autopilots and guidance and navigation algorithms.

In our experience at Honeywell's Systems and Research Center over the last five years, we have seen these methods applied to numerous aerospace control system analysis and design problems: involving the F-15 STOL DEMO vehicle, Space Shuttle, and NASP to name just a few. In current control design applications the SSV is the key ingredient that insures needed multivariable robustness properties of the feedback control laws.

The Structured Singular Value concept was in its infancy when this ONR-sponsored program began more than eight years ago. The rapid growth of the concept during those early years led to the ONR/Honeywell Workshop, a three day event in October of 1984. Featured speakers at that Workshop were (among others) Gunter Stein, John Doyle, Bruce Francis. The notes from the presentations of these three speakers constituted the official set of Workshop Notes. Those Notes constituted Volume 1 of this final report.

It is satisfying to see a theoretical concept grow and find practical applications as successfully as the SSV concept has over the last 8 years. This ONR sponsored program has provided valuable support for the basic mathematical research effort. The successful collaboration we have had between top-notch mathematics and control experts in academics and industry is not easy to keep alive without some source of government research funds.

When it came time to write the final report, there were several possibilities that came to mind concerning what to write about. Ideas about documenting the history of the SSV development, a serious study of robust, practical control law design, a (finally) comprehensible development of the SSV theory with a survey and comparison of accumulated results, etc.

After several unsuccessful attempts to write on several of those sensible and useful topics, it became clear that the final report was going to be -- more mathematical theory (this is a research contract, after all).

Technical Report Summary

The underlying mathematical μ-theory was and still remains an evolving bundle of concepts and techniques, often discussed and written about by the investigators, never completely formalized. This report is primarily devoted to a presentation of a relatively small piece of the overall SSV theory and the simplest part at that: the diagonal, unrepeated complex block problem. One would have thought that after eight years of research the simplest part of the theory would have been worked out to everyone's satisfaction. In fact, for the most simple, nontrivial μ-problem (2x2 diagonal unrepeated blocks) it has.

After the notation in section 1 of the report, we present to everyone's satisfaction (or, at least the author's satisfaction) the 2x2 diagonal-block μ-theory. If all we wanted was a statement and proof of the result we could have made that section much shorter. The greater effort put into that simple problem was part of a (not very well concealed) plan to bring invariant theory into the picture.

To see the invariants, and the role they play in the problem, the theory must be polished to a very fine resolution. Every parameter of the problem must be accounted for. After going through the theory in this new way we started to see things a little differently. That was the plan, *because our secret goal was to solve, at last,* the four-block diagonal μ-problem by a method that could be implemented efficiently on a computer. That simple problem was the first μ-problem on the hit-list after Doyle proved his remarkable theorem for the three-block case in 1982. The author tried his hand at it and failed. It remained unsolved throughout the duration of this contract. Clearly, we were missing something, so a fresh approach looked like a good idea.

We solved the four-block diagonal μ-problem. We did it while trying to use invariant theory. We developed a reasonable algorithm, implemented and tested it on a computer, and it seemed to work (with no other way to find the exact answer, how do we know if it works? -- it did agree with Packard's lower bound to within a couple of percent). Plots of the results are shown at the end of Section 3. Then we started to write the details of the proofs. The theory started to change (but, remarkably, the algorithm did not). The theory below (still in flux) is the current version.

Many of the results presented here were known earlier, but there are several ideas that seem original. A partial list of highlights is:

1) Theorem 3.1 (with Last Minute Remarks at the end of section 4):

> This theorem is a canonical form description of the diagonal, non-repeated block μ-problem. Though not made truly canonical until the report was almost finished, in its incomplete form it laid a foundation for the geometric analysis. In its complete form it allows immediate classification of the 3-block μ-problem within a space depending (generically) on 8 real parameters (the four-block mu problem appears to depend on 23 real parameters).

2) Theorem 3.3

This theorem led to the development of the computational algorithm that can now be used to compute $\mu(M)$ exactly for the 4-diagonal block problem. It provides a set of polynomial equations, equations that are very easy to understand and work with (if not always easy to solve), and shows how the solutions to that set of equation give solutions to the μ-problem.

3) Constructive algorithm 3.1

There is a general result from classical elimination theory stating that a system of n+1 homogeneous polynomial equations in n unknowns can be solved -- that method can (in principle) be used to solve the μ-problem given the result of Theorem 3.3. The constructive algorithm we use though not formalized yet, is a form of elimination specially tailored for the polynomial system at hand. There might be better methods available, but this one was readily understood, easy to code, numerically robust, etc.

4) Loose End 4.1

This rambling discussion describes the very recent efforts toward applying invariant theory to improve our understanding of structured singular values, with an eye toward efficient computational methods. In the absence of any new or practically useful results to show for our efforts so far, we have focused attention on explaining why we approached the problem from this viewpoint and what we can realistically hope to achieve. The final observation made at the deadline, Observation 4.4, explains how we intend to proceed.

The treatment here is far from a comprehensive study of structured singular values. For those who want a broader reference, the best we know is the 1988 AFOSR Report "Robust Control of Multivariable and Large Scale Systems," by Andy Packard and John Doyle.

The parameter spaces used here are the same as those used in past studies. The new idea in this report is to look at the problem from a global perspective, to find all the points that look like they could be $\mu(M)$, and determine how to pick the right one. Previous studies quickly eliminated all of the bogus μ-like points, but they were forced to work locally to do so (to the credit of their inventors, the local techniques have led to some good global results).

The characterization of the set S^{sing} in Theorem 3.2 and the correspondence between those points and the real points on a singular complex variety given in Theorem 3.3 turns out to be natural (at least the author still thinks so). The problem with this global perspective is the large number of complex points on that variety that eat up computer time when one has to use in order to carry them along. For the three and four block problems the dimensions work out so that the global evaluation is feasible, but already for dimension 5 we are not sure if this approach is still practical.

The final section on loose ends discusses and sometimes improves upon the shortcomings of results in earlier sections. A major part of that section is devoted to an informal discussion of what we are trying to do with invariants. There is no telling when we will succeed in solving the problems stated in that section, but the success we have had implementing code to solve these problems confirms that these problems are solvable.

In short, we solved the problem we set out to solve but probably not much more. In the process we have come to a better understanding of the complexity of the μ-problem, and we have some less than certain approaches to extending the computable results out to 6 or 8 blocks or so. The picture revealed in going through the theory is fascinating, with inevitable ties to algebraic geometry and classical invariant theory. We hope the reader will find something of interest in the presentation.

Programmatic Overview

This program was people working on theory, talking to each other, and writing papers for publication.

This report was written by the Honeywell Principal Investigator, but the ideas presented (at least the good ones) are primarily the results of consultants who worked on the program and of colleagues at Honeywell Systems and Research Center. Names of those whose efforts contributed directly this program are:

Consultants, Visitors, ...

1) John Doyle - Star of the team, SSV Inventor
2) Andy Packard
3) Allen Tannenbaum
4) Pramod Khargonekar
5) Mike Safonov
6) Jim Freudenberg
7) Bruce Francis

Honeywell Personnel and Expatriots

1) Gunter Stein
2) Gary Hartmann
3) Mike Barrett
4) Dale Enns
5) Jim Krause
6) Kathryn Lenz
7) Chester Chu
8) Mike Elgersma
9) Dan Bugajski
10) Joe Wall
11) Art Harvey
12) Blaise Morton - Honeywell PI

Apologies to those who played a role and are not listed here. It is hard to remember all the contributors to an eight-year old program. It is also difficult to keep track of who gets credit for everything. You guys know what you did.

References

During the course of the program there were many publications partially funded by this contract, or for which there may have been overlap of activities (e.g. AFOSR). No complete list of these papers has been maintained, on the next page is a partial list.

PARTIAL LIST OF PAPERS FUNDED BY CONTRACT

[1] J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainty," IEE Proceedings, Part D, V 129, No. 6, Nov., 1982.

[2] J. C. Doyle, J. E. Wall, and G. Stein, "Performance and Robustness Analysis for Structured Uncertainty," 1982 Conference on Decision and Control.

[3] J. C. Doyle, "Robust Stability with Structured Uncertainty," 1985 Mathematical Theory of Networks and Systems, Stockholm.

[4] J. C. Doyle, &Lecture~Notes&, 1984 ONR/Honeywell Workshop on Advances in Multivariable Control, October 8-10, 1984, Minneapolis.

[5] J. C. Doyle, "Synthesis of Robust Controllers and Filters," 1983 Conference on Decision and Control.

[6] J. C. Doyle, "Structured Uncertainty and Robust Multivariable Control Design," 1985 IFAC Workshop on Model Error Concepts and Compensation, Boston.

[7] B. G. Morton and R. M. McAfoos, "A Mu-test for Robustness Analysis of a Real-parameter Variation Problem," 1985 ACC, Boston.

[8] B. G. Morton, "New Applications of Mu to Real-Parameter Variation Problems," 1985 CDC, Ft. Lauderdale.

[9] J. C. Doyle and C-C. Chu, "Matrix Interpolation and &H sub inf& Performance Bounds," 1985 ACC, Boston.

[10] C-C. Chu and J. C. Doyle, "The General Distance Problem in &H sub inf& Synthesis," 1985 Conference on Decision and Control, Ft. Lauderdale.

[11] J. C. Doyle, "Structured Uncertainty in Control System Design," 1985 CDC, Ft. Lauderdale.

[12] Doyle, J. C., Robust Control of Multivariable and Large Scale Systems, Honeywell Final Report on Contract No. F49620-82-C-0090 to AFOSR, for work during period July 1982 through June 1984, submitted May 1985.

[13] Doyle, J. C. and C. C. Chu, Robust Control of Multivariable and Large Scale Systems, Honeywell Final Report on Contract No. F49620-84-C-0088 to AFOSR, for work during period July 1984 through October 1985, submitted March 1986.

[14] Packard, A., and J. C. Doyle, Robust Control of Multivariable and Large Scale Systems, Honeywell Final Report on Contract No. F49620-86-C-0001 to AFOSR, for work during period October 1985 through March 1988, submitted March 1988.

[15] Doyle, J. C., K. Lenz and A. Packard, "Design Examples Using &mu& Synthesis: Space Shuttle Lateral Axis FCS During Reentry," Proceedings of the CDC, December 1986, pp. 2218-2223.

[16] Doyle, J. C., R. S. Smith, and D. F. Enns, "Control of Plants with Input Saturation Nonlinearities," Proceedings of the American Control Conference, June 1987, pp. 1034-1039.

[17] Krause, J. M., "Stability Margins with Real Parameter Uncertainty: Test Data Implications," Proceedings of the American Control Conference, 1989.

[18] Doyle, J. C., K. Glover, P. P. Khargonekar, and B. A. Francis, "State-Space Solutions to Standard &H sub 2& and &H sub inf& Control Problems," IEEE Transactions on Automatic Control, Vol 34, No. 8, August 1989, pp. 831-847.

[19] Lenz, K., H. Ozbay, A. Tannenbaum, J. Turi, and B. Morton, "Robust Design for a Flexible Beam Using a Distributed-Parameter &H sub inf& Controller," Submitted to Proceedings of CDC in September, 1989.

[20] Foias, C., A. Tannenbaum, D. Enns, T. Georgiou, M. Jackson, and B. Schipper, "On the Nonlinear Mixed Sensitivity Problem," Preprint.

[21] Balas, G. and J. Doyle, "On the Caltech Experimental Large Space Structure," 1988 ACC.

[22] Packard, A., and J. Doyle, "Quadratic Stability with real and Complex Perturbations," Preprint

[23] Smith, R., J. Doyle, M. Morari, A. Skjellum, "A Case Study Using μ: Laboratory Process Control Problem," Preprint

[24] Packard, A. and J. Doyle, "Structured Singular Value with Repeated Scalar Blocks,"

There are many other publications that were either completely or partially funded by this contract. We are willing to provide a more complete list if desired.

For another novel approach to the μ problem (not partially funded by this contract), there is a very fine piece of recent work co-authored by one of our co-investigators:

Bercovici, H., C. Foias, and A. Tannenbaum, "Structured Interpolation Theory," Preprint.

One of the interesting results in this paper is a new proof of Doyle's three-block theorem -- that the upper-bound is equal to μ in the case of a three-block Δ structure.

1. Notation

Let M be a complex $n \times n$ matrix and denote by Δ the set of block diagonal matrices:

$$\Delta = \left\{ \begin{bmatrix} \Delta_1 & 0 & 0 & 0 \\ 0 & \Delta_2 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \Delta_k \end{bmatrix} \right\} \tag{1}$$

where each Δ_j is a complex $m_j \times m_j$ matrix , $m_1 + \cdots + m_k = n$. In this structure one or more blocks on the diagonal may be repeated -- an important special case is the set of scalar matrices, denoted $\delta\ Id_n$, that commutes with all $n \times n$ matrices.

In this note the scalar matrices are the only structure considered with repeated blocks. This restriction to the non-repeated case allows a simplification of notation: when we want to indicate a specific structure we affix the values $m_1, \cdots, m_k$ as superscripts to Δ (e.g. $\Delta^{1,1,1,1}$ is the set of diagonal 4×4 matrices). When the specific structure is not important we simply use the symbol Δ and place the burden on the reader to remember that a set of positive integers m_j summing to n is tacitly assumed.

Associated with Δ is the set U_Δ of block-diagonal unitary (i.e. $U^*U = Id$) matrices that are contained in Δ. For example, when Δ is the set of diagonal matrices we have:

$$\Delta_j = e^{i\theta_j} \tag{2}$$

where θ_j is a real number.

We are interested in solving the following maximization problem: find

$$\max\{\ \bar{\rho}\ (X\ M) \mid X\ \varepsilon\ U_\Delta\ \} \tag{3}$$

where $\bar{\rho}$ is the spectral radius function. The solution to this maximization problem was shown by Doyle to be the structured singular value of M, denoted $\mu(M)$, for the structure Δ.

There are two special structures for which μ can be identified with standard, important functions. The spectral radius of M is the function $\mu(M)$ associated with the structure $\delta\ Id_n$, and the maximum singular value of M is the function $\mu(M)$ associated with the structure Δ^n. These two special cases of the function μ are extreme in that $\mu(M)$ for any other structure lies between those two values.

Another special case arises when Δ is the set of diagonal $n \times n$ matrices: then $k = n$ and each m_j is 1. It is this special case we will discuss below.

2. The Simplest Non-Trivial Example

We will consider the general (but unrepeated-block) structure in a later section. The anxious reader can skip this section, but it is probably easier to read this section first. Here we see how the general approach works in the simplest non-trivial example: $\mu(M)$ associated with the structure $\Delta^{1,1}$.

Let v denote a general complex 2-vector ($v \varepsilon C^2$)

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} . \tag{4}$$

The set U_Δ consists of 2x2 matrices X of the form:

$$X = \begin{bmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{bmatrix} . \tag{5}$$

For each X in U_Δ let us examine the eigenvalue problem for the product matrix XM. If the complex scalar λ is an eigenvalue of XM there must be a nonzero v such that equation 6 holds:

$$XM\ v = \lambda\ v . \tag{6}$$

To solve the maximization problem in equation 3 we need to determine the largest $|\lambda|$ that can arise for any X in U_Δ and nonzero v satisfying equation 6.

In the past, two types of approach have been applied to this problem:

Approach 1: Start at $X = Id_2$ and solve the eigenvalue problem for XM = M. Next, for j = 1,2, consider X_j obtained by incrementing θ_j by a small step-size $d\theta_j$. Solve the eigenvalue problems for X_jM, compute max $|\lambda|_j$ in both cases, and increment θ_1, θ_2 by taking a small step in the direction of greatest first-order increase in $|\lambda|$. Iterate until a maximum is found.

Approach 2: Using the parametric representation of U_Δ given in equation 5, form the product matrix XM symbolically as a function of θ_1, θ_2 and the entries of M. Write down the analytic expression for the two eigenvalues of XM as functions of these same variables. Differentiate with respect to θ_1, θ_2 the expressions for the magnitudes of the two eigenvalues at each point and set these partial derivatives equal to zero to obtain a set of equations to be solved for θ_1, θ_2 in order to determine the critical values of $|\lambda|$. Solve all these equations for θ_1, θ_2 and substitute the solutions back into the eigenvalue formulas. The largest $|\lambda|$ arising from the finite set of points considered should be the desired maximum.

These two approaches can be used but neither one provides a satisfactory solution to the general problem. The problem with the first approach is the possibility of local maxima less than the true maximum. Examples have been constructed where multiple local maxima do exist. Consequently, gradient search techniques do not have guaranteed convergence properties.

The second approach has the problem that it does not generalize to a tractable algorithm for problems involving larger matrices.

The underlying weakness of both approaches is the reliance on analytic properties of the eigenvalue problem depending on the X-parameter in equation 6. In our approach we solve the maximization problem of equation 3 without making direct use of equation 6. We can sidestep the eigenvalue problem because it gives more information than we really need. The key question is how big we can make $|\lambda|$ before equation 6 has no solution for any X in U_Δ. It is possible to reformulate the problem in such a way that the the vector v, the phase of λ, and the matrix X do not play a direct role. In its new form the problem can be solved.

For the structure $\Delta^{1,1}$ the reformulated problem is stated in terms of a pair of 2x2 Hermitian matrices $H^1(r)$, $H^2(r)$ depending on a real parameter r. These two matrices are defined in dyadic terms by:

$$H^1(r) = \begin{bmatrix} \overline{m}_{11} \\ \overline{m}_{12} \end{bmatrix} [m_{11} \; m_{12}] - \begin{bmatrix} 1 \\ 0 \end{bmatrix} r \, [1 \; 0] \, , \quad H^2(r) = \begin{bmatrix} \overline{m}_{21} \\ \overline{m}_{22} \end{bmatrix} [m_{21} \; m_{22}] - \begin{bmatrix} 0 \\ 1 \end{bmatrix} r \, [0 \; 1] \tag{7}$$

where m_{ij} are the components of the matrix M. The following lemma motivates the construction of $H^i(r)$.

Lemma 1: There exist X ε U_Δ, nonzero v, and complex λ such that $|\lambda| = r$ satisfying equation 6 if and only if there is a nonzero vector y ε C^2 such that

$$y^* H^1(r)\, y = 0 \qquad \text{and} \qquad y^* H^2(r)\, y = 0 \, . \tag{8}$$

Proof of Lemma 1: First suppose a nonzero y satisfies equation 8. The components y_1, y_2 of y satisfy:

$$|m_{11}y_1 + m_{12}y_2|^2 = r\, |y_1|^2 \quad |m_{21}y_1 + m_{22}y_2|^2 = r\, |y_2|^2 \tag{9}$$

Taking square roots of both sides of both equations and setting $\lambda = \sqrt{r}$, we find that equation 9 is equivalent to the existence of real numbers θ_1, θ_2 such that

$$\begin{bmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} . \tag{10}$$

The expression in equation 10 is the component form of equation 6.

Conversely, suppose equation 10 is satisfied. It is easy to see that the equations in 9 are satisfied if r is set equal to $|\lambda|^2$. Then equation 8 is also satisfied. Lemma 1 is proved.

In view of lemma 1 we may restate the maximization problem of equation 3 as follows: determine the largest real number r such that a nonzero vector y exists satisfying equation 8. This reformulation leads immediately to

Question 1: Given a pair of Hermitian forms H^1 and H^2, when does there exist a nonzero vector y such that $y^* H^1 y = 0$ and $y^* H^2 y = 0$ are satisfied simultaneously?

We will answer question 1 below in the statement of lemma 2. First we make two simple observations.

Recall that the eigenvalues of a Hermitian matrix are real. We call a Hermitian matrix definite if all its eigenvalues are nonzero and have the same sign (i.e. all are positive or all are negative).

Observation 1: If either form H^1 or H^2 is definite then no solution vector y exists.

Returning to the definitions of the matrices $H^i(r)$ in equation 7, however, we see that both matrices are indefinite by construction. Each has one positive and one negative eigenvalue if no row of M is zero. So observation 1 provides no useful information, but it does lead directly to

Observation 2: If there is a real vector $t = [t_1, t_2]$ such that $H(t) = t_1 H^1 + t_2 H^2$ is definite then no solution vector y exists.

This second observation will be useful. In particular, if r is chosen sufficiently large, we see that $H^1(r) + H^2(r)$ will be negative definite. Thus observation 2 can be used to place an upper bound on the size of $\mu(M)$. Better yet, it leads to

Lemma 2: If H(t) is indefinite for all real vectors t, then there exists a nonzero vector y such that

$$y^* H^1 y = 0 \quad \text{and} \quad y^* H^2 y = 0$$

are satisfied simultaneously.

Proof of Lemma 2: It suffices to consider the one-parameter family of matrices

$$F = \{t H^1 + H^2 \mid t \in R\} . \tag{11}$$

A Hermitian matrix has real eigenvalues; its determinant is the product of those eigenvalues. It follows that a 2×2 Hermitian matrix is definite if and only if its determinant is positive.

Let us suppose that every matrix in F is indefinite. We conclude:

$$\text{for all real } t \quad \det(t H^1 + H^2) \le 0 \tag{12}$$

There are now two cases.

Case 1: $\det(H^1) = 0$ -- then the function of t on the left-hand side of inequality 12 is an affine, non-positive function that therefore must be constant (i.e. independent of t). These conditions can arise for a pair of Hermitian, 2×2 matrices only if H^1 is a scalar multiple of H^2 , in which case the conclusion of lemma 2 follows (take any nonzero y such that $y^* H^2 y = 0$.

Case 2: $\det(H^1)$ nonzero -- change the basis of C^2 (if necessary) so that the matrices of H^1 and H^2 are:

$$H^1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} , \quad H^2 = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{13}$$

for some real numbers h_{11}, h_{22} and complex numbers $h_{12} = \bar{h}_{21}$.

In these coordinates the inequality 12 is equivalent to

$$(h_{11} + h_{22})^2 - 4h_{12}h_{21} \le 0 \tag{14}$$

Also, any vector y satisfying $y^* H^1 y = 0$ has coordinates $y_1 = ae^{i\theta_1}$, $y_2 = ae^{i\theta_2}$ for some real numbers θ_1, θ_2 and a. We want to find a nonzero vector of this form that also satisfies $y^* H^2 y = 0$. We might as well set a = 1.

The proof will be complete if we can find two real numbers θ_1 and θ_2 such that:

$$\begin{aligned} 0 &= [e^{-i\theta_1}, e^{-i\theta_2}] \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} e^{i\theta_1} \\ e^{i\theta_2} \end{bmatrix} \\ &= h_{11} + h_{22} + 2\mathrm{Re}(e^{i(\theta_2 - \theta_1)} h_{12}) \end{aligned} \tag{15}$$

But equation (15) can be solved if and only if the inequality in 14 holds. Lemma 2 is proved.

Remark 1: As a consequence of Lemma 2 and the preceding observations, we have proved the following:

Let H^1, H^2 be a pair of indefinite 2×2 Hermitian matrices. Then one of two exclusive alternatives holds:

Alternative 1: The matrix $t_1 H^1 + t_2 H^2$ is definite for some pair of real numbers t_1, t_2.

Alternative 2: There is a nonzero vector y such that $y^* H^i y = 0$ for i = 1,2.

Remark 2: The proof of Lemma 2 does not generalize to higher dimensions, nor does the conclusion if n is bigger than 3. For higher dimensions we need to use a more sophisticated approach as described in Section 3.

Let us conclude the analysis of the $\Delta^{1,1}$ mu_problem.

For the matrices $H^1(r)$ and $H^2(r)$ defined in equation 8, define the set F(r):

$$F(r) = \{t\, H^1(r) + H^2(r) \mid t \in R\} \tag{16}$$

We have already observed that $H^1(r) + H^2(r)$ is definite for r sufficiently large. Define R_0 to be the infimum of the set of R such that for all $r \ge R$ there exists a pair of real numbers t_1, t_2 such that $t_1 H^1(r) + t_2 H^2(r)$ is definite. R_0 is finite and non-negative. By this construction, any real combination of $H^1(R_0)$ and $H^2(R_0)$ must be indefinite, so alternative 2 of remark 1 above must hold. Furthermore, for any $r > R_0$ alternative 1 must hold (hence 2 cannot): it follows that R_0 is the solution to the optimization problem, hence equal to $\mu(M)^2$ (recall that in equation 10 we set $r = |\lambda|^2$ and $\mu(M)$ is the maximum $|\lambda|$).

How do we compute R_0? First, we know from what we have already seen that there is a nonzero vector y such that $y^* H^i(R_0)\, y = 0$. Consequently, there is a unitary matrix U_0 (not necessarily in U_Δ) such that for all $t \in R$:

$$U^* \left(t\, H^1(R_0) + H^2(R_0)\right) U = \begin{bmatrix} 0 & t\, h^1{}_{12} + h^2{}_{12} \\ t\, h^1{}_{21} + h^2{}_{21} & t\, h^1{}_{22} + h^2{}_{22} \end{bmatrix} . \tag{17}$$

Fortunately, we can say even more about the family $F(R_0)$

Lemma 3: Let $H^1(r)$, $H^2(r)$, $F(r)$ and R_0 be defined as above. For some real t^0 the matrix

$$H^0 = t^0 H^1(R_0) + H^2(R_0) \tag{18}$$

is semi-definite, with at least one zero eigenvalue.

Proof of Lemma 3: For any $\varepsilon > 0$ the family $F(R_0 + \varepsilon)$ contains a definite matrix. There is no difficulty in finding a compact set B of real numbers (B depends on $H^1(r)$ and $H^2(r)$ for r between R_0 and $R_0 + 1$) such that for any ε between 0 and 1 a number t^ε in B satisfies $t^\varepsilon H^1(R^0 + \varepsilon) + H^2(R_0 + \varepsilon)$ is definite.

Take a sequence $\{\varepsilon_i\}$ decreasing to 0, and select an accompanying sequence $\{t^{\varepsilon_i}\}$ in B such that the corresponding matrix is definite. Let t^0 be any limit point of the sequence t^{ε_i}, and define $H^0 = t^0 H^1(R_0) + H^2(R_0)$ as in equation 18. Every open set of 2×2 Hermitian matrices containing H^0 also contains a definite matrix, yet H^0 is not definite. The eigenvalues are continuous functions, so H^0 must be semi-definite; because it is not definite it has at least one zero eigenvalue.

Lemma 3 is proved.

With lemma 3 in hand, let us reexamine the significance of equation 17. We now know that when t^0 of lemma 3 is substituted for t in 17 the determinant vanishes. In fact, the function

$$p(t) = \det(t\, H^1(R_0) + H^2(R_0)) \tag{19}$$

must have a zero of second order at $t = t^0$. This means that the derivative polynomial $\frac{dp}{dt}(t)$ is also zero at t_0. Summarizing this last discussion, we make

Observation 3: At $r = R_0$ three things happen:

1) all forms in $F(R_0)$ are indefinite
2) for some t^0 the form $t^0 H^1(R_0) + H^2(R_0)$ is semidefinite
3) when the form is semidefinite, the determinant vanishes at t^0 to second order in t.

From these facts we see that R_0 can be computed by the following algorithm.

Step 1: Compute the coefficients of the polynomial function $p_1(t,r)$ of two variables defined by

$$p_1(t,r) = \det(t\, H^1(r) + H^2(r)) \tag{20}$$

Step 2: Compute the coefficients of

$$p_2(t,r) = \frac{\partial p_1}{\partial t}(t,r) \tag{21}$$

Step 3: Compute the coefficients of the polynomial function q(r) obtained by eliminating t from p_1 and p_2.

Step 4: Find the largest positive real root of the polynomial q(r) (if no positive real root to q(r) exists, take 0 for the answer).

Claim: The answer obtained in the fourth step is R_0, the square of the desired value $\mu(M)$. There are two cases to consider in verifying this last claim.

Case 1: $m_{11} = 0 = m_{12}$ -- then the only nonzero entry in $H^1(r)$ is -r in the 1,1 spot and $\det(H^1(r))$ is zero for all r. Consequently, the polynomial $p_1(t,r)$ is at most first-order in t so $p_2(t,r)$ is independent of t, and R_* satisfies $p_2(t,R_*) = 0$ if and only if $|m_{22}|^2 = R_*$. Then the 2,2 entries of $H^1(R_*)$ and $H^2(R_*)$ are both zero, hence alternative 2 of remark 2 is satisfied at $r = R_*$. If $r > R_*$, the 2,2 entry of H^2 is negative and so, for large enough t, the determinant of $t\,H^1 + H^2$ is positive. Thus alternative 1 holds for $r > R_*$, hence $R_* = R_0 = \mu(M)^2$.

Case 2: m_{11} or m_{12} nonzero-- then $\det(H^1(r))$ is negative for any $r > 0$ so $p_1(t,r)$ is quadratic in t. If R_* is any positive real root of the polynomial q(r) then $p_1(t,R_*)$ is of the form

$$p_1(t,R_*) = -k_*^2\,(t - t_*)^2$$

where the leading coefficient is negative because $\det(H^1(R_*))$ is negative. It follows that at R_* the determinant is non-positive for all t, hence $t\,H^1(R_*) + H^2(R_*)$ is indefinite for all t. Thus R_0 cannot be less that R_*, the largest positive solution of q(r). On the other hand, R_0 is also be a root of q(r), hence $R_* = R_0 = \mu(M)^2$ as claimed.

Remark 3: For this low dimensional example the answer can be computed without going through the formal procedure just outlined. When the size of the problem gets bigger, however, the complexity of the computations becomes much greater and a more systematic approach (e.g. elimination theory) is required. Though the details are different for the higher dimensional problem discussed in the next section, the algorithm to compute $\mu(M)$ is basically the same. A numerical example illustrates the four-step approach for the structure $\Delta^{1,1}$.

Numerical Example: Let M be the 2×2 matrix

$$M = \begin{bmatrix} 1.0 & 10.0 \\ 0.1 & i \end{bmatrix} \tag{22}$$

The two matrices $H^1(r)$ and $H^2(r)$ are:

$$H^1(r) = \begin{bmatrix} (1.0 - r) & 10.0 \\ 10.0 & 100.0 \end{bmatrix} \quad H^2(r) = \begin{bmatrix} 0.01 & 0.1i \\ -0.1i & (1.0 - r) \end{bmatrix} \tag{23}$$

Step 1 for example: The polynomial $p_1(t,r)$ is:

$$p_1(t,r) = \det(t\, H^1(r) + H^2(r)) \tag{24}$$
$$= -100.0\, r\, t^2 + ((1 - r)^2 + 1)\, t - 0.01\, r$$

Step 2 for example: The polynomial $p_2(t,r)$ is:

$$p_2(t,r) = \frac{\partial p_1}{\partial t}(t,r) \tag{25}$$
$$= -200.0\, r\, t + (r^2 - 2.0\, r + 2.0)$$

Step 3 for example: Setting equation 25 equal to zero and solving for t

$$t = \frac{(r^2 - 2.0\, r + 2.0)}{200.0\, r} \tag{26}$$

Obtain q(r) by substituting equation 26 into equation 24 and clearing fractions:

$$q(r) = (r^2 - 2.0\, r + 2.0)^2 - 4.0\, r^2 \tag{27}$$
$$= (r^2 + 2.0)\, (r^2 - 4.0r - 2.0)$$

Step 4 for example: The largest root of q(r) is:

$$R_0 = 2.0 + \sqrt{2.0} \tag{28}$$

Finally

$$\mu(M) = \sqrt{R_0} = \sqrt{2.0 + \sqrt{2.0}} \tag{29}$$

3. The Theory for Diagonal Matrices

In this section we analyze the case of diagonal, unrepeated blocks.

Let M be a complex $n \times n$ matrix and denote by $\Delta^{1,\cdots,1}$ the set of $n \times n$ diagonal matrices:

$$\Delta^{1,\cdots,1} = \left\{ \begin{bmatrix} \delta_1 & 0 & 0 & 0 \\ 0 & \delta_2 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \delta_n \end{bmatrix} \right\} \tag{3.1}$$

where each δ_j is a complex number. Here we do not allow repetition of any of the δ_j symbols on the diagonal.

Recall that U_Δ is the set of $n \times n$ unitary matrices in $\Delta^{1,\cdots,1}$: i.e. for $j = 1, \cdots, n$ $|\delta_j| = 1$. For each X in U_Δ we consider the eigenvalue problem for the product matrix XM. If the complex scalar λ is an eigenvalue of XM there must be a nonzero v such that equation 3.2 holds:

$$XM\, v = \lambda\, v\, . \tag{3.2}$$

As in section 2, we need to determine the largest $|\lambda|$ that can arise for any X in U_Δ and nonzero v satisfying equation 3.2.

We now reformulate the problem in terms of n Hermitian $n \times n$ matrices $H^1(r), \cdots, H^n(r)$ depending on a real parameter r. These matrices are defined in dyadic terms by:

$$H^1(r) = \frac{1}{r} \begin{bmatrix} m_{11}^* \\ m_{12}^* \\ \cdots \\ m_{1n}^* \end{bmatrix} [m_{11}\ m_{12}\ \cdots\ m_{1n}] - \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} [1\ 0\ \cdots\ 0]\, , \tag{3.3}$$

$$\cdots$$

$$H^n(r) = \frac{1}{r} \begin{bmatrix} m_{n1}^* \\ m_{n2}^* \\ \cdots \\ m_{nn}^* \end{bmatrix} [m_{n1}\ m_{n2}\ \cdots\ m_{nn}] - \begin{bmatrix} 0 \\ \cdots \\ 0 \\ 1 \end{bmatrix} [0\ 0\ \cdots\ 1]\, .$$

Lemma 3.1 Let R_0 be the largest real number for which there exists a nonzero vector y satisfying:

$$y^* H^j(r)\, y = 0 \quad j = 1, \cdots, n. \tag{3.4}$$

Then $\mu(M) = \sqrt{R_0}$.

Proof of lemma 3.1: This simple computation is analogous to the one for the special case (see Lemma 1 of Section 2).

Remark 3.1 The matrices $H^j(r)$ defined here differ from those in Section 2 by a factor of $\frac{1}{r}$. The idea is to normalize the H^j matrices so that the negative eigenvalue is -1 for any r. The disadvantage of this change is that the $H^j(r)$ matrices are no longer defined at r = 0. Because we are primarily interested in those matrices for which $\mu(M)$ is greater than zero we accept this disadvantage for the sake of normalization.

The matrices $H^j(r)$ are generically rank 2, but there are special conditions where the rank drops to 0.

Definition 3.1 We will call the matrix M degenerate if some row is zero off the main diagonal (i.e. for some i and all j not equal to i, $m_{ij} = 0$).

It is convenient to exclude the degenerate matrices from the general analysis below so we take care of them now.

Lemma 3.2: Suppose the matrix M is degenerate, and that row i is zero away from the diagonal. Then $\mu(M)$ is the larger of the two numbers:

1) $|m_{ii}|$ or
2) $\mu(M_i)$

where M_i is the $(n-1) \times (n-1)$ matrix obtained by deleting the i^{th} row and column of M.

Proof of Lemma 3.2: Let M, v and λ satisfy equation equation 3.2. Two possibilities arise as follows:

1) if the vector v has v_i nonzero then $|\lambda| = |m_{ii}|$,

2) if the vector v has $v_i = 0$ then the n-1 vector obtained from v by deleting v_i satisfies 3.2 for the matrix M_i and the same λ.

We know that $\mu(M)$ is at least as large as $|m_{ii}|$ because m_{ii} is an eigenvalue of M and $\mu(M) \geq \bar{\rho}(M)$. The only way that $\mu(M)$ could be larger is if some v with $v_i = 0$ satisfies equation 3.2 with some λ of magnitude larger than $|m_{ii}|$. In that case $\mu(M)$ is equal to $\mu(M_i)$.

Lemma 3.2 is proved.

If the matrix M_i obtained from M is still degenerate, lemma 3.2 can be applied repeatedly. Repeated application will eventually lead to the computation of μ for a non-degenerate matrix or, if M is diagonal, eliminate the μ computation completely (for diagonal M, $\mu(M) = \bar{\rho}(M)$).

For the rest of this section we assume M is non-degenerate.

All the matrices $H^j(r)$ are rank 2. Let us assume there is a nonzero vector y satisfying equation 3.4. Then there is a unitary transformation U such that the 1,1 entries of the matrices $U^* H^j(r)\, U$ are 0. In theorem 1 below we characterize the structure of the matrices $U^* H^j(r)\, U$.

Theorem 1 (Canonical Form): Suppose there is a nonzero vector y satisfying equation 3.4. Then there is a unitary matrix U such that for each j the matrix $H^j(r)$ is in the following form:

$$U^* H^j(r)\, U = \begin{bmatrix} 0 & \bar{k}^j\,(\alpha^{j*} - \beta^{j*}) \\ k^j\,(\alpha^j - \beta^j) & \alpha^j\alpha^{j*} - \beta^j\beta^{j*} \end{bmatrix} \tag{3.5}$$

The n complex numbers k^j, the n (n-1)-vectors α^j and β^j in 3.5 are related to U and M by the equations:

$$U = \begin{bmatrix} k^1 & \beta^{1*} \\ \cdots & \cdots \\ k^n & \beta^{n*} \end{bmatrix} \qquad \frac{e^{i\Theta}\, M\, U}{\sqrt{r}} = \begin{bmatrix} k^1 & \alpha^{1*} \\ \cdots & \cdots \\ k^n & \alpha^{n*} \end{bmatrix} \tag{3.6}$$

for some real diagonal $n \times n$ matrix Θ.

<u>Proof of Theorem 1:</u> Constructive -- by assumption there is a unit length vector k such that

$$e^{i\Theta}\, \frac{M}{\sqrt{r}} \begin{bmatrix} k^1 \\ \cdots \\ k^n \end{bmatrix} = \begin{bmatrix} k^1 \\ \cdots \\ k^n \end{bmatrix} \tag{3.7}$$

The unit vector k can be expanded (nonuniquely) by n complex vectors β^j of size (n-1) to a $n \times n$ unitary matrix U:

$$U = \begin{bmatrix} k^1 & \beta^{1*} \\ \cdots & \cdots \\ k^n & \beta^{n*} \end{bmatrix} \tag{3.8}$$

The vectors α^j of size n-1 are then given by the following equation:

$$\frac{e^{i\Theta}\, M\, U}{\sqrt{r}} = \begin{bmatrix} k^1 & \alpha^{1*} \\ \cdots & \cdots \\ k^n & \alpha^{n*} \end{bmatrix} \tag{3.9}$$

Equations 3.8 and 3.9 are the identities in equation 3.6. From here it is an easy task to verify 3.5: the j^{th} row of the right hand side of 3.9 can be multiplied on the left by its conjugate transpose to get

$$\begin{bmatrix} |k^j|^2 & \bar{k}^j\, \alpha^{j*} \\ k^j\, \alpha^j & \alpha^j\alpha^{j*} \end{bmatrix}$$

which is one part of of 3.5 (the matrix $e^{i\Theta}$ drops out). Similarly the j^{th} row of the right hand side of 3.8 can be left-multiplied by its conjugate transpose to get:

$$\begin{bmatrix} |k^j|^2 & \bar{k}^j \beta^{j*} \\ k^j \beta^j & \beta^j \beta^{j*} \end{bmatrix}$$

When these last two matrices are subtracted the result is the right hand side of equation 3.5.

Theorem 1 is proved.

Remark 3.2 Except for special cases of M the vector k is unique up to phase. The vector k can be identified with the vector y that was assumed to exist in the hypotheses. There was some freedom in the selection of the parameters β^j -- all other solutions are obtained by multiplying U on the right by a general matrix of the form:

$$T_U = \begin{bmatrix} 1 & 0 \\ 0 & T_{n-1} \end{bmatrix} \tag{3.10}$$

where T_{n-1} is a unitary matrix of size $n-1 \times n-1$. The α^j vectors are uniquely defined once the other parameters are fixed.

The representation of the $H^j(r)$ matrices given in Theorem 1 applies to any value $r = |\lambda|^2$ for which the μ-eigenvalue equation 3.2 has a solution. We are interested in characterizing extremal solutions of 3.2. More precisely, we want to determine those values R_k that are local maxima of those values of r for which equation 3.2 can be solved.

Before embarking on our extremal set analysis there are two topics to discuss. First, a summary of our goals for the rest of this section.

We are headed for a theorem that characterizes the solution to the structured singular value problem for a generic set of matrices. What we mean by generic is the second topic in this digression, for now we concentrate on describing the general approach. The first step is to identify a real-algebraic set S contained in C^n that contains the extremal point we seek. The set S has a simple definition but a complicated structure. At a generic point S has a neighborhood diffeomorphic to R^n, but there are exceptional points comprising the singular set of S where this is not true. It is in the singular set of S that our answer lies.

Our final answer will be a set of polynomial equations that can be solved to determine every point in the singular set of S. This set of polynomial equations is derived from a set of singularity relations on the Jacobian of the defining equations of S that must hold at a point y_0 if it is in the exceptional set. The canonical form in Theorem 1 is used to construct the polynomial equations for the singular set from the singularity relations. Elimination theory is used to generate a polynomial in the single variable r. The largest real solution of this polynomial is (generically) the square of $\mu(M)$.

As mentioned at the beginning of this digression, the argument below will apply only to generic matrices M. For our work here, an argument will be said to apply generically if it is true for an open (standard Euclidean norm topology), dense subset of the space of complex $n \times n$ matrices. The precise conditions for which arguments apply will be stated in each case. The results presented here were derived only recently and we did not have time to look for more general proofs. Presumably, the special cases will be worked out later and a more comprehensive treatment might be found.

Returning to the extremal problem we introduce more notation.

Let S_r be the following set:

$$S_r = \{y \,\varepsilon\, C^n \mid y^* y = 1 \,,\, y^* H^j(r)\, y = 0 \;\; j = 1, \cdots, n\} \,. \tag{3.11}$$

We want to find the largest value of r such that S_r is nonempty. Define the set S to be the union of S_r over all $r > 0$. Suppose we have a curve y(r) mapping r in some interval $[r_0, r_1]$ into S such that $y(r) \,\varepsilon\, S_r$. We suppose that the mapping is differentiable at all points where it is defined. What can be said about a value R_k of r beyond which this curve cannot be continued?

At all points r for which the curve y(r) is defined differentiably, the tangent vector $\frac{dy}{dr}$ must satisfy:

$$\frac{dy^*}{dr}\, y + y^* \frac{dy}{dr} = 0 \qquad \frac{dy^*}{dr}\, H^j(r)\, y + y^*\, H^j(r)\, \frac{dy}{dr} + y^*\, \frac{dH^j(r)}{dr}\, y = 0 \tag{3.12}$$

$$j = 1, \;\cdots\;, n \,.$$

Altogether, this set of equations imposes n+1 real-linear conditions on the tangent vector $\frac{dy}{dr}$ that lives in a 2n-dimensional real vector space. The only way that there could be a problem in finding a solution for a given value y(r) is if the equations are overdetermined and inconsistent. Theorem 2 characterizes the local maximum points R_k.

Theorem 2:

1) Suppose y_0 is a point in S_{r_0}. If the space of vectors $w \,\varepsilon\, C^n$ satisfying

$$w^*\, y_0 + y_0^*\, w = 0 \qquad w^*\, H^j(r_0)\, y_0 + y_0^*\, H^j(r_0)\, w + y_0^*\, \frac{dH^j(r_0)}{dr}\, y_0 = 0 \tag{3.13}$$

$$j = 1, \;\cdots\;, n$$

is exactly (n-1)-dimensional then r_0 is not a local maximum point.

2) The local maximum points R_k are contained in the singular set S^{sing} defined as follows: S^{sing} is the set of those values of r for which there exists a set of real numbers $\{t_1 \;\cdots\; ,t_n\}$ not all zero and $y_k \,\varepsilon\, S_r$ such that:

$$\sum_{j=1}^{n} t_j\, H^j(r)\, y_k = 0 \quad \text{and} \quad t_j\, k_j \text{ nonzero for some j.} \tag{3.14}$$

Proof of Theorem 2:

1) The existence of an (n-1)-dimensional space of solutions w of equation 3.13 at y_0 implies a local manifold structure for S (locally n-dimensional) near y_0. A curve through y_0 can be found for values of r in the interval $[r_0, r_0 + \eta]$ for some small, positive η that stays in S. It

follows that r_0 is not a local maximum point.

2) The alternative to case 1 can arise only for values $r = R_k$ for which there is a point y_0 in C^n where the Jacobian of the mapping from C^n to R^{n+1} defining S_{R_k} is not full rank. Let us suppose that there is a linear relation among the n+1 gradient functions. We use the canonical form coordinates provided by Theorem 1 for a general point in S. In these coordinates, the point y_0 and the general tangent vector w are:

$$y_0 = \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad w = \begin{bmatrix} a_1 + ib_1 \\ a_2 + ib_2 \\ \cdots \\ a_n + ib_n \end{bmatrix} \tag{3.15}$$

The first relation $w^* y_0 + y_0^* w = 0$ is, in this notation:

$$2a_1 = (a_1 - ib_1) + (a_1 + ib_1) = 0 \tag{3.16}$$

or, in other words, w_1 is pure imaginary. This condition is the infinitesimal form of the requirement that $\|y(r)\|$ have a constant value: the curve stays in S^{2n-1}, the unit 2n-1 sphere in C^n. The space S is contained in $S^{2n-1} \times R$ where the last R factor parametrizes the r-values.

There are n more linear relations imposed on w: for each j we have

$$[(a_1 - ib_1) \quad (a_2 - ib_2) \quad \cdots \quad (a_n - ib_n)] \begin{bmatrix} 0 \\ k^j(\alpha^j - \beta^j) \end{bmatrix} + \tag{3.17}$$

$$[0 \quad \overline{(k^j} \, (\alpha^{j*} - \beta^{j*}))] \begin{bmatrix} a_1 + ib_1 \\ a_2 + ib_2 \\ \cdots \\ a_n + ib_n \end{bmatrix} + [1 \; 0 \; \cdots \; 0] \, U^* \frac{dH^j(R_k)}{dr} \, U \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = 0 \, .$$

The expression $\dfrac{dH^j(R_k)}{dr}$ is computed in the original coordinates of equation 3.2. We have chosen our coordinates so that the derivatives of U and U^* do not appear: the matrices U are held fixed (at their values for the canonical form at R_k) as the parameter r is varied continuously about R_k. For now we treat the last summand on the left hand side of 3.17 as a general (indeterminate) vector. To simplify notation define:

$$Z^j = -\, [1 \; 0 \; \cdots \; 0] \, \frac{d\, U^* \, H^j(R_k)\, U}{dr} \begin{bmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} . \tag{3.18}$$

Then (3.17) simplifies to:

$$2\,\mathrm{Re}\left[\begin{bmatrix}\bar{k}^1\,(\alpha^{1*}-\beta^{1*})\\ \cdots \\ \bar{k}^n\,(\alpha^{n*}-\beta^{n*})\end{bmatrix}\begin{bmatrix}a_2+ib_2\\ \cdots \\ a_n+ib_n\end{bmatrix}\right]=\begin{bmatrix}Z^1\\ \cdots \\ Z^n\end{bmatrix}. \tag{3.19}$$

If the rows of the complex coefficient matrix $[\bar{k}^j\,(\alpha^{j*}-\beta^{j*})]$ are linearly independent over the real numbers, then for any Z there exists a set of real numbers a_j and b_j solving equation 3.19. Conversely, if the rows are linearly dependent over the reals then equation 3.19 can be solved only if Z lies in the kernel of the real row vector that annihilates the complex coefficient matrix. The local manifold structure of S is guaranteed only at those points where the columns are independent over the reals. This leads us to define:

Singularity Conditions: let r be in S and suppose y_0 is in S_r. Choose a transformation U so that $U^* H^j(r)\,U$ is in the canonical form of Theorem 1. Let Z be the real vector in C^n given by 3.18. If r is in S^{sing} then there must be some set $\{t_1, \cdots, t_n\}$ of nonzero real numbers such that:

$$[t_1 \;\cdots\; t_n]\begin{bmatrix}\bar{k}^1\,(\alpha^{1*}-\beta^{1*})\\ \cdots \\ \bar{k}^n\,(\alpha^{n*}-\beta^{n*})\end{bmatrix}=0 \quad t_j\bar{k}_j \text{ nonzero for some j.} \tag{3.20}$$

Equation 3.14 is an equivalent form of the singularity conditions in the original coordinates (without the U^* U terms, and conjugate-transposed).

Theorem 2 is proved.

Remark 3.3: The conditions imposed in 3.16 are independent of those imposed in 3.19. Geometrically this makes sense -- the relations in 3.19 are the infinitesimal form of the relations $y^* H^j(r)\,y = 0$ that are invariant under multiplication by a complex scalar and so define a relation on the projective space of lines in C^n. Equation 3.16, on the other hand, is the infinitesimal form of the normalization of vectors by their length. Any (nonzero) projective-space solution of the equations $y^* H^j(r)\,y = 0$ can be normalized to have unit length, so it is consistent that the infinitesimal constraints should be independent.

Remark 3.4: As mentioned in Remark 3.2 the canonical form of Theorem 1 is not unique. Because that form is used in the definition of S^{sing}, we should verify that the set S^{sing} is well defined, independent of choices. A different choice of form would result in multiplying the complex coefficient matrix $\bar{k}^j\,(\alpha^{j*}-\beta^{j*})$ on the right by a unitary matrix T_{n-1} as in equation 10 (actually, the phase of the matrix T_{n-1} may be shifted by a a uniform factor due to nonuniqueness of the vector k). Transformations of this type cannot change the singularity status of this matrix -- the singularity test depends on the existence of a real vector that annihilates the complex coefficient matrix when applied on the left; while the nonuniqueness is represented by multiplication with a nonsingular matrix on the right. Thus S^{sing} is well defined, independent of choices made in the construction.

The space S^{sing} has a direct analytic tie to the original μ-problem. Consider a value r for which S_r is nonempty. From the equations 3.6 it is easy to verify that:

$$\left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right]U = \begin{bmatrix} 0 & \alpha^{1*} - \beta^{1*} \\ \cdots & \cdots \\ 0 & \alpha^{n*} - \beta^{n*} \end{bmatrix} . \tag{3.21}$$

It is clear from this expression that $\sqrt{r}$ is a valid eigenvalue for the equation 3.2 because the rank of

$$\left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right]$$

is less than n. For such a matrix there is always a vector w^* such that

$$w^* \left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right] = 0 . \tag{3.22}$$

The vector w^* is the left eigenvector for the problem 3.2. If r is in S^{sing}, however, then by 3.20 and 3.21 there is a set of real numbers $\{t_1, \cdots, t_n\}$ such that $t_j \bar{k}^j$ is nonzero for some j and

$$[t_1 \bar{k}^1, \cdots, t_n \bar{k}^n] \left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right]U = 0. \tag{3.23}$$

Recall that k is the right eigenvector for M from Theorem 1. An alternate characterization of S^{sing} is as follows:

<u>Alternate Singularity Characterization:</u> The point r is in S^{sing} if there is a real $n \times n$ matrix Θ such that the right eigenvector k and left eigenvector w^* for $e^{i\Theta}M$ have components with conjugate phase, i.e. $w^{j*} k^j$ is real for $j = 1, \cdots, n$.

So far we have worked at reformulating the problem but we have not made much progress toward solving it. Before proceeding to the solution we pause to introduce a generic condition on the class of M.

GENERIC CONDITION: We assume that for all real r the rank of

$$\left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right] \tag{3.24}$$

for any real matrix Θ is never less than n-1.

Remark 3.5: Another way of stating the generic condition is to require that the matrix $e^{i\Theta}M$ is free of repeated eigenvalues for any real Θ. Experience with numerical aspects of the standard eigenvalue problem suggests that this restriction on the problem might simplify analysis. The possibility of nontrivial Jordan cells is eliminated if repeated eigenvalues are not allowed. The non-repeated root condition holds for an open, dense subset of M and so is a generic condition. It will be assumed to hold for M in the analysis that follows.

It is worthwhile to see what the generic condition implies about the canonical form.

Observation 3.1 The generic condition is exactly the condition that the matrix:

$$\left[\frac{e^{i\Theta}M}{\sqrt{r}} - Id_n\right] = \begin{bmatrix} 0 & \alpha^{1*} - \beta^{1*} \\ \cdots & \cdots \\ 0 & \alpha^{n*} - \beta^{n*} \end{bmatrix} U^* .$$

(see equation 3.21) is rank n-1. Equivalently, the $n \times (n-1)$ matrix $[\alpha^{j*} - \beta^{j*}]$ is full rank. The direction of the vector $[t_1 \bar{k}^1, \cdots, t_n \bar{k}^n]$ that appears in equation 3.23 as a left eigenvector is therefore uniquely determined (up to a complex scalar factor) by the matrix $[\alpha^{j*} - \beta^{j*}]$.

Theorem 3: Let M be a generic matrix, and consider the polynomial function in the real vector $t = \{t_1, \cdots, t_n\}$ and $\frac{1}{r}$:

$$p(t_1, \cdots, t_n, \frac{1}{r}) = \det(t_1 H^1(r) + \cdots + t_n H^n(r)) \tag{3.25}$$

If r_0 is in S^{sing} then the function p is a polynomial that is not identically zero. Furthermore, there exists some point $\{t_1^0, \cdots, t_n^0\}$ such that

$$p(t_1^0, \cdots, t_n^0, \frac{1}{r_0}) = 0, \tag{3.26a}$$

$$\frac{\partial p}{\partial t_j}(t_1^0, \cdots, t_n^0, \frac{1}{r_0}) = 0 \qquad j = 1, \cdots, n. \tag{3.26b}$$

Furthermore, from 3.26a and 3.26b there can be constructed a polynomial $g(\frac{1}{r})$ such that every real number r_0 in S^{sing} is a solution of

$$g(\frac{1}{r_0}) = 0 \tag{3.27}$$

Conversely, if r_0 is a real root of the polynomial g and if equations 3.26a and 3.26b are satisfied nonvacuously by some real set of values $(t_1^0, \cdots, t_n^0)$ then r_0 is in S.

Finally, if S^{sing} is nonempty, $\mu(M) = \sqrt{R_{max}}$ where R_{max} is the largest real number in S^{sing}. If S^{sing} is empty, $\mu(M) = 0$.

<u>Proof of Theorem 3:</u>

First we want to show that if r_0 is in S^{sing} there is a nonzero real vector t^0 such that p and $\frac{\partial p}{\partial t_j}$ vanish at $(t^0, \frac{1}{r_0})$. By Theorem 1 there is a unitary U such that

$$U^* H^j(r_0) U = \begin{bmatrix} 0 & \bar{k}^j (\alpha^{j*} - \beta^{j*}) \\ k^j (\alpha^j - \beta^j) & \alpha^j\alpha^{j*} - \beta^j\beta^{j*} \end{bmatrix} \tag{3.28}$$

By Theorem 2 there is a real nonzero vector t^0 such that

$$[t_1^0 \cdots t_n^0] \begin{bmatrix} \bar{k}^1 (\alpha^{1*} - \beta^{1*}) \\ \\ \bar{k}^n (\alpha^{n*} - \beta^{n*}) \end{bmatrix} = 0 . \tag{3.29}$$

Fixing the parameter r_0, define H(t) by

$$H(t) = t_1 H^1(r_0) + \cdots + t_n H^n(r_0) \tag{3.30}$$

Then

$$U^* H(t) U = \begin{bmatrix} 0 & \sum_{j=1}^{n} t_j \bar{k}^j (\alpha^{j*} - \beta^{j*}) \\ \sum_{j=1}^{n} t_j k^j (\alpha^j - \beta^j) & \sum_{j=1}^{n} t_j (\alpha^j\alpha^{j*} - \beta^j\beta^{j*}) \end{bmatrix} . \tag{3.31}$$

The 1,1 entry in 3.31 vanishes identically, the first row and column vanish at $t = t^0$. Expanding the determinant about the first row it is clear that the polynomial

$$p(t, \frac{1}{r_0}) = \det(H(t)) = \det(U^* H(t) U) \tag{3.32}$$

vanishes to first order at $t = t^0$ (saying that H(t) vanishes to first order at t^0 is another way of saying that $H(t^0) = 0$ and $\frac{\partial H}{\partial t^j}(t^0) = 0$). It is not difficult to show (using the generic assumption on M and equation 3.31) that the unique (up to scale) t^0 satisfies the polynomial equations 3.26a and 3.26b nonvacuously (see below). We have verified that each r_0 in S^{sing} gives rise to a nontrivial solution of equations 3.26a and 3.26b.

We now want to show the converse. Suppose r_0 is a value of r such that the real vector t^0 satisfies equations 3.26a and 3.26b nonvacuously (nonvacuously means that at least one nonzero t^0-monomial is multiplied by a nonzero coefficient of p).

Compute the $n-1 \times n-1$ minor-polynomials of H(t) with respect to some row -- say the first row. Let $f_i(t)$ denote the i^{th} minor polynomial. Define the vector polynomial function F(t):

$$F(t) = \begin{bmatrix} f_1(t) \\ \cdots \\ f_n(t) \end{bmatrix} . \tag{3.33}$$

From $\det(H(t^0)) = 0$ it follows that

$$H(t_0)F(t_0) = \begin{bmatrix} \det(H(t^0)) \\ 0 \\ \cdots \\ 0 \end{bmatrix} = 0 \tag{3.34}$$

From the condition that the determinant of H(t) vanishes at t^0 to first order, we know that

$$\frac{\partial}{\partial t_j}\Big[F^*(t)\ H(t)\ F(t)\Big] \Big|_{t=t^0} = \tag{3.35}$$

$$\frac{\partial F^*}{\partial t_j}(t^0)\ H(t^0)\ F(t^0) + F^*(t^0)\ \frac{\partial H}{\partial t_j}(t^0)\ F(t^0) + F^*(t^0)\ H(t^0)\ \frac{\partial F}{\partial t_j}(t^0) = 0 .$$

By equation 3.34 and its conjugate transpose, the first and third terms of the sum in 3.35 vanish. Evaluating the middle term, we find:

$$F^*(t^0)\ H^j\ F(t^0) = 0 . \tag{3.36}$$

which is exactly the form of the homogeneous system of equations defining S. The argument is not complete, however, because all the first order minors could vanish at t^0 -- then $F(t^0)$ is the zero vector and equation 3.36 reduces to a vacuous assertion.

In case all the minors of H(t) vanish at t_0 , take repeated partial derivatives of the function F(t) with respect to an appropriate set of t_i, creating a sequence of vector polynomials $F_0(t) = F(t), F_1(t), F_2(t), \cdots$ that all vanish to first order at t^0 until, finally, $F_n(t)$ vanishes at t^0 but no longer to first order. This process works because of the assumption that t_0 satisfies the equations 3.26a and 3.26b (it would not be true otherwise). Suppose $\frac{\partial F_n}{\partial t_j}(t^0)$ is nonzero. Then

$$0 = \frac{\partial H}{\partial t_j}(t^0)\, F_n(t^0) + H(t^0)\frac{\partial F_n}{\partial t_j}(t^0) \tag{3.37}$$

so $H(t)\frac{\partial F_n}{\partial t_j}(t)$ vanishes at t^0 and $\frac{\partial F_n}{\partial t_j}\,(t_0)$ is nonzero.

Repeat the computations of equations 3.35 and 3.36 with $\frac{\partial F_n}{\partial t_j}$ in place of F.

We conclude that r_0 is in S.

To complete the proof of Theorem 3.3 we explain how the polynomial g is constructed. The system of equations 3.26a and 3.26b is a set of n+1 simultaneous equations in the n+1 unknown variables $t_1, \cdots, t_n, \frac{1}{r}$. In fact, these polynomials are not independent because of the Euler identity:

$$n\, p(t_1^0, \cdots, t_n^0) = \sum_{j=1}^{n} t_j \frac{\partial p}{\partial t_j}(t_1^0, \cdots, t_n^0, \frac{1}{r_0}) \tag{3.34}$$

Consequently, any solution of equations 3.26b will automatically satisfy equation 3.26a, so we really have only n equations to solve. The polynomials are homogeneous in the t-variable, however. For each solution (t,r) with t nonzero there is a 1-parameter family of solutions $(\lambda t, r)$ for all real λ. The solution for each fixed r is a real projective variety in the t-space, so the system of polynomials depends on n-1 independent affine parameters.

The general technique for solving this type of problem is elimination theory, as described in [Van der Waerden]. The resulting polynomials can have very large degree, however, and it might be better to take advantage of the structure of the polynomial system. The technique used to test the theory on the 3 and 4 block cases is shown below as Constructive Algorithm 3.1.

All the constructions are now complete. Denote by V the the set of real solutions r_j of the resultant polynomial $g(\frac{1}{r})$. In the first part of the proof we showed that S^{sing} is a subset of V. In the second part we showed that V is a subset of S. The largest value R_{max} of S lies in S^{sing}, therefore V is a finite set subset of S containing R_{max}. For $r > R_{max}$ we have S_r is empty, but $S_{R_{max}}$ is nonempty. From Lemma 3.1, if S^{sing} is nonempty then $\mu(M) = \sqrt{R_{max}}$. If S is empty then R_{max} is not defined and $\mu(M) = 0$.

Theorem 3 is proved.

Constructive Algorithm 3.1: In the proof of Theorem 3.3 we referred to a constructive algorithm that takes advantage of the structure of the polynomial system. We illustrate the algorithm here with the example from the 3-block case.

The most general form of the determinant polynomial in the 3-block scalar case is:

$$\det(H(t)) = c_{210}\, t_1^2 t_2 + c_{201}\, t_1^2 t_3 + c_{120}\, t_1 t_2^2 + \tag{3.35}$$

$$c_{111}\, t_1t_2t_3 + c_{102}\, t_1t_3^2 + c_{021}\, t_2^2t_3 + c_{012}\, t_2t_3^2$$

The coefficients c_{ijk} are cubic polynomials in $\frac{1}{r}$. The three partial-derivative functions are easily computed:

$$\frac{\partial H}{\partial t_1} = 2c_{210}\, t_1t_2 + 2c_{201}\, t_1t_3 + c_{120}\, t_2^2 + c_{111}\, t_2t_3 + c_{102}t_3^2\,, \tag{3.36a}$$

$$\frac{\partial H}{\partial t_2} = c_{210}\, t_1^2 + 2c_{120}\, t_1t_2 + c_{111}\, t_1t_3 + 2c_{021}\, t_2t_3 + c_{012}\, t_3^2 \tag{3.36b}$$

$$\frac{\partial H}{\partial t_3} = c_{201}\, t_1^2 + c_{111}\, t_1t_2 + 2c_{102}t_1t_3 + c_{021}\, t_2^2 + 2c_{012}t_2t_3\,. \tag{3.36c}$$

Our goal is to determine a relation on the coefficients that holds whenever the system gives a nontrivial point in S^{sing}. The approach is to construct a 12×12 matrix out of the c_{ijk} such that the determinant vanishes whenever the r that they depend upon lies in S^{sing}.

The procedure for constructing the matrix is as follows. For each of the three derivative polynomials, form the product with four separate monomials:

$$T_1 = \{t_1^2,\, t_1t_2,\, t_1t_3,\, t_2t_3\} \tag{3.37a}$$

$$T_2 = \{t_1t_2,\, t_1t_3,\, t_t^2,\, t_2t_3\} \tag{3.37b}$$

$$T_3 = \{t_1t_2,\, t_1t_3,\, t_2t_3,\, t_3^2\} \tag{3.37c}$$

The result is 12 polynomials that are generated by the 12 monomials:

$$B = \{t_1^3t_2,\, t_1^3t_3,\, t_1^2t_1^2,\, t_1^2t_2t_3,\, t_1^2t_3^2,\, t_1t_2^3,$$
$$t_1t_2^2t_3,\, t_1t_2t_3^2,\, t_1t_3^3,\, t_2^3t_3,\, t_2^2t_3^2,\, t_2t_3^3\} \tag{3.38}$$

Form the coefficient matrix for the 12 polynomials. If there is a solution to the original three equations for which the function H(t) does not vanish identically (at least 2 of the three values t_1,t_2,t_3 must be nonzero) then the determinant of the 12×12 coefficient matrix will vanish. Conversely, if the determinant vanishes then there is some solution of the original three equations with the property that at least 2 of the three values t_1,t_2,t_3 are nonzero.

To generate the polynomial g of Theorem 3.3, substitute the appropriate polynomial expression into the c_{ijk} functions that appear in the expression for the determinant of the coefficient matrix.

We solved the four-block case by a similar, explicit scheme. The result was (for one example) a 68×68 matrix that generically had rank 62 (we also tried a degenerate (definition 3.1) example that generically had rank 49).

Remark 3.7: When testing the algorithm we did not explicitly evaluate the polynomial g -- there was no need. Instead, we wrote a computer program that generated, for a given value of r, the coefficient matrix corresponding to that value. The value of r was then allowed to vary, by small increments, over a specified range of values. At each value of r an appropriate singular value of the coefficient matrix was computed using the LINPACK program dsvd. A plot was then drawn for visual inspection. Some sample plots are shown in the accompanying figures on the next page.

The plots on the next page show the results when the exact 4-block μ computation algorithm is run on an example. The example matrix is the sample test matrix M in the family of matrices in [Robust Control of Multivariable and Large Scale Systems, Final technical Report to AFOSR, Contract No. F49620-86-C-0001, March 23, 1988 by Andy Packard], page 39 with parameters defined on page 40. The interesting aspect of this example is that the upper and lower bounds do not agree, so no technique for finding the exact value of μ was known until the constructive algorithm 3.1 was developed.

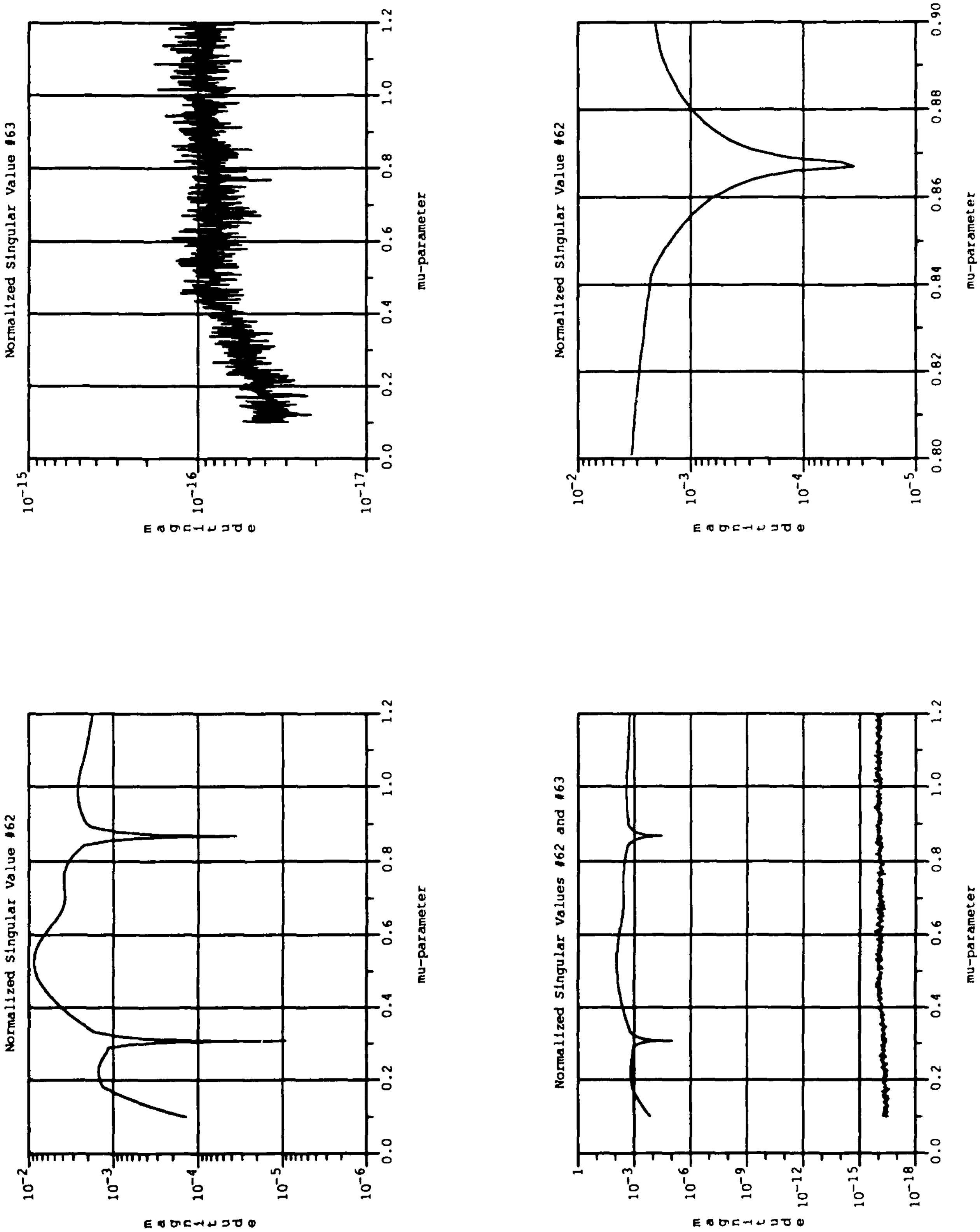
Using the constructive algorithm in the 4-block case, we generated a 1-parameter family of 68×68 matrices that had entries linear in the c_J variables. There were 6 relations in the generator set, so the rank of the matrix was generically 62. The horizontal axis of the plot is the so-called mu-parameter. The largest mu-parameter value on the plot where the the matrix looses rank is $\sqrt{R_{max}} = \mu(M)$. Note that there appear to be only two nonzero mu-parameters where the matrix looses rank -- this observation is significant because it means that the majority of the 62 roots of the associated polynomial are probably complex, leaving what could be a low order invariant factor that contains the sought point R_{max} (if more roots were real we would expect to see many dips in the plot).

The second plot on the page shows the value of the singular value #63. Note that it is more than 10 orders of magnitude smaller than the average value of singular value #62. This gap indicates a clear boundary where the matrix becomes rank-deficient. The third plot shows the two singular values plotted to the (logarithmic) scale for direct comparison of magnitudes.

Finally, the last plot shows a blowup of the region of interest near the numerical zero of singular value #62. For comparison, the lower bound program mup gave a lower bound of 0.8723 and an upper bound of 1.000 for $\mu(M)$.

The constructive algorithm required roughly 6 SPARC second to evaluate the 68 singular values at each mu-parameter value on the plot. For a typical plot consisting of 200 points, the time required to evaluate a 4-block mu value exactly is just over 10 minutes. This is more than 100 times slower than mup requires to provide the upper and lower bounds.

07/31/90 16:18, EXACT 4x4 MU COMPUTATION



4. Observations and Loose Ends

Theorem 3.3 in the previous section is an important breal. ,rough for the μ-theory. It provides a theoretical algorithm for computing $\mu(M)$, and it led to the functional algorithm we now have on the computer that can solve the four-block diagonal problem.

That theorem alone, however, does not provide a practical solution for the diagonal μ-problem of arbitrary size. The computational method we have implemented does not generalize easily to arbitrary dimensions. There are related theoretical developments still in progress, however, that could improve this situation. This section is mainly concerned with other research in progress that, together with Theorem 3.3, could bring us to a more complete understanding of the μ-problem.

One point worth mentioning right away is we are not sure whether S^{sing} is a good set of points to work with. We have worked only a few examples in the case of 2, 3 and 4 scalar blocks, and little is known in general about the structure of S and how S^{sing} sits inside it. It seems likely that the theory would improve immensely if the global structure of some examples were worked out in detail. That is the long-term goal of the research described in this section.

Observation 4.1: Much of the theory presented in Section 3 was developed in a different framework by Andy Packard. In his Ph. D. thesis, he derived a set of polynomial conditions that he used to define μ-values [Packard 1]. There is a clear correspondence between his polynomials and those in Theorem 3.2. The iterative lower-bound algorithm that Andy developed [Packard 2] is based on a decomposition that makes use of a stationarity condition of an associated gradient function.

Observation 4.2: There could be a better statement of the result in Theorem 3.3 more closely related to both the computational methods and the geometric structure. For robust computation we choose to compute the points of S^{sing} by means other than polynomial-root finding. As mentioned in Remark 3.7 at the end of section 3, the earliest numerical test of the theory on 3 and 4 block examples did not rely on finding the roots of the polynomial g (we did not even compute the polynomial g).

Observation 4.3: The reader should note that the methods used to obtain the polynomial conditions in the n-block case in Section 3 are different from those used in the 2-block case presented in Section 2. In Remark 2.1 an alternative criterion for pairs of indefinite 2×2 matrices was stated, and our initial approach to the general theory was based on a generalization of that alternative condition. We state the (incorrect) generalization here.

Incorrect General Statement 4.1: Let H^j $j = 1, \cdots, n$ be a set of indefinite $n \times n$ Hermitian forms. Define $Z(H^j)$ to be the zero set of H^j viewed as a quadratic form on C^n:

$$Z(H^j) = \{y \in C^n \mid y^* H^j y = 0\} . \tag{4.1}$$

Then one of two alternatives holds:

1) There exist real numbers $t_1, \cdots, t_n$ such that $\sum_{j=1}^{n} t_j H^j$ is definite

2) The intersection of $Z(H_j)$ over all j contains a nonzero point.

This statement is true for n=2 but not true for n≥4. The incorrectness of this general statement for $n \geq 4$ is linked to the gap between $\mu(M)$ and $\inf\{\bar{\sigma}(D\ M\ D^{-1}) \mid D\ \varepsilon\ GL_\Delta\}$ (the upper bound found by Doyle in the original paper [Doyle] in which μ was defined). The approach in Section 2 might be extended to higher n as a way to relate the upper bound to μ. Finding an exact bound on the size of the gap, even for the case $n = 4$, would be another significant breakthrough.

Loose End 4.1: So far we have not directly referred to invariants. In fact, much of this approach was motivated by the idea that invariant theory could uncover some practical results for the numerical solution of the problem. The function det(H(t)) is one type of invariant function, there are other invariants as well.

Only recently have we started looking at other invariants associated with the μ-problem. The 2-block case is trivial, the 3-block case is already hard but it appears tractable. Our status in these two cases is presented below.

The two-block case: When $n = 2$ the polynomial det(H(t)) is:

$$\det(H(t_1, t_2)) = c_{20}\, t_1^2 + c_{11}\, t_1 t_2 + c_{02}\, t_2^2\ . \tag{4.2}$$

The condition that the gradient of det(H(t)) vanish at t^0 is:

$$\frac{\partial(\det(H(t^0))}{\partial t_1} = 2c_{20}\, t_1^0 + c_{11}\, t_2^0 = 0\ ,\ \frac{\det(H(t^0))}{\partial t_2} = c_{11}\, t_1^0 + 2c_{02}\, t_2^0 = 0\ . \tag{4.3}$$

These two equations can be satisfied for a nonzero t^0 only if the discriminant function Δ defined by

$$\Delta(c_{20}, c_{11}, c_{02}) = 4\, c_{20}c_{02} - c_{11}^2 = 0\ . \tag{4.4}$$

Recall that the coefficients c_{ij} are polynomial functions of $\frac{1}{r}$. The polynomial $g(\frac{1}{r})$ referred to in Theorem 3.3 is (functionally equivalent to) the polynomial Δ when it is evaluated as a function of r.

The three-block case: When $n = 3$ the polynomial det(H(t)) is:

$$\det(H(t)) = c_{210}\, t_1^2 t_1 + c_{201}\, t_1^2 t_3 + c_{120}\, t_1 t_2^2 + \tag{4.5}$$

$$c_{111}\, t_1 t_2 t_3 + c_{102}\, t_1 t_3^2 + c_{021}\, t_2^2 t_3 + c_{012}\, t_2 t_3^2\ .$$

The problem is: find a set of polynomial functions of the coefficients c_{ijk}, analogous to the discriminant function Δ for the two-block case, such that the vanishing of those functions is equivalent to the condition that the gradient of g vanish at a point where at least 2 of the components t_j^0 are nonzero (we may assume that two t_j^0 are nonzero because of the non-vacuous assumption in Theorem 3.3).

It is possible that the trace of H(t) and the sum of second principal minors will help solve the problem (there is a discriminant for cubic polynomials in two variables that include these quantities).

The technique developed by mathematicians to solve this type of problem is invariant theory, the solution may be found (in principle) by analysis of the induced representations of the general linear group GL(n) acting on degree-r homogeneous polynomials in t. All the basic invariants can be determined by classical methods (see [Weyl]).

Though this theoretical answer is appealing, it is never clear in practice what invariants to compute. The recipe indicated in [Weyl] is to use the so-called symbolic method, a technique that can be used to generate all the invariants of a given fixed degree. But then, he goes on to say:

> "Great as this accomplishment is, one ought to point out, however, that the method is far from reducing the construction of a finite integrity basis for form invariants to the same for vector invariants. For the number of symbolic vector arguments $u^1, \cdots, u^v$ we have to introduce [during the invariant construction process] is dependent on the degree of J(u) [the invariant function], and we must have an unlimited supply of such symbols at our disposal when we are to take into account invariants J of all possible degrees."
>
> [Weyl, p. 244]

In addition, there are general problems for which the full solution, even if computable, might be impractical to implement in a computer program.

For the system of polynomials in Theorem 3.3, however, there is some hope for a satisfactory solution in low dimensions. At least in the three and four block cases the invariant computations are of a small enough size that they can be performed by hand.

For example, one of the simplest low-degree invariants, the Hessian

$$X(t) = \det\left[\frac{\partial^2 \det(H(t))}{\partial t_i \partial t_j}\right] \tag{4.6}$$

(a covariant of weight 2, see [Weyl, p. 240]) provides a large set of nontrivial invariants to work with. Note that

$$X(t) = \sum_{i_1, \cdots, i_n} f_{i_1, \cdots, i_n}(c_J)\, t_1^{i_1} \cdots t_n^{i_n} \tag{4.7}$$

is a polynomial in t homogeneous of degree n(n-2) in t with coefficient functions $f_{i_1, \cdots, i_n}(c_J)$ that are homogeneous of degree n in the c_J (J is a multi-index, e.g. J = 2022 for the coefficient c_{2022} when n = 4). The summands in equation 4.7 are assumed to have been symmetrized, with the summation running over all indices $i_1, \cdots, i_n$ such that $\sum_{j=1}^{n} i_n = n(n-2)$.

Our objective is to find relations among sets of invariants that hold whenever the value of r they depend on lies in S^{sing} (see Theorem 3.2). (Note: an expression for the Hessian matrix and its determinant has been computed directly from the expression in equation 3.31. There might be symmetries but we have no definite results yet).

Now we already knew that there is a single polynomial $g(\frac{1}{r})$ that these coefficients had to satisfy: it can be generated by Sylvester's determinants in elimination theory (see [Van der

Waerden]). The problem with Sylvester determinants is the degree of the resultant polynomial grows in size rapidly as n increases: the resultant degree is (generically) the product of the degrees of the individual polynomials. For the system of polynomials in Theorem 3.3 there are n equations each of degree n-1 -- the growth in degree is asymptotically n^n.

One good reason for believing there should be invariant factors for Sylvester's determinant polynomial is the existence of Constructive Algorithm 3.1. It is just a matter of identifying the irreducible representation of the general linear group associated with that construction and we will have an invariant relation to work with.

Very recently, we checked to see what the Hessian polynomials look like for the case of n=3. In all, there are 10 polynomials of degree 3 in the 7 nonvanishing coefficients of the polynomial det(H(t)) in equation 4.5. Two of the 10 quantities (generalized discriminants) are presented in equation 4.10.

$$c_{111}c_{210}c_{201} - c_{120}c_{201}^2 - c_{102}c_{210}^2 \tag{4.10a}$$

$$c_{210}\,(8\,c_{201}c_{021} - 4\,c_{210}c_{012} - 4\,c_{120}c_{102} + 2\,c_{111}^2) \tag{4.10b}$$

Equation 4.10a is the coefficient of t_1^3 and 4.10b is the coefficient of $t_1^2t_2$ in the polynomial X(t). There are two quantities similar to 4.10a and five others similar to 4.10b (the six relations of type 4.10b seem to be subdivided into two groups of three according to orientation). There is one other quantity of a type not shown -- that one arises from the coefficient of $t_1t_2t_3$ in X(t).

We have not solved the problem, but we can state precisely what we hope for: THE BEST POSSIBLE RESULT.

THE BEST POSSIBLE RESULT: For the three-block problem, the best possible result would be to find a small set of low-degree invariants, similar in form to those in equation 4.10a and 4.10b, that generate the full ring of invariant functions. It would then follow that the Sylvester determinant polynomial, being an invariant polynomial, would lie in the polynomial ideal generated by those polynomials. For the examples in equation 4.10 each coefficient c_J is a polynomial of degree 3 in the parameter $\frac{1}{r}$, that would therefore lead to generators of order no greater than 9.

If such a result could be found then the Sylvester's determinant could be factored into primitive invariants and then, to compute $\mu(M)$, one needs only find all the real roots of the factor polynomials, call the largest one R_{max}, then $\mu(M) = \sqrt{R_{max}}$.

We emphasize that we have not achieved the best possible result, nor do we feel totally confident that we can solve it. On the other hand we have good reason to expect some progress. Some hope seems justified because the problem originated from a system of quadratic forms $y^{j*}\,H^j\,y^j = 0$, and problems involving sets of quadratic forms have been solved in the past [Bromwich]. We have only seen results for one parameter families, however (our problem is exactly the one Bromwich treats, but for n-parameter families and lower rank forms).

It is known that there is a finite set of polynomial generators and the classical theory provides exhaustive methods for finding such a set. Even so, the problem of demonstrating an explicit set of generators is usually not easy. Fortunately, there is another possibility:

AN EASIER RESULT OF EQUAL PRACTICAL VALUE: There is a deterministic process that can be used to tell in finite time whether a given set of invariants is good enough to solve the μ-problem. For our application it would be sufficient, and no loss in terms of practical value, if we could find any set of low order invariants that factoize the Sylvester polynomial. This restriction allows us to reformulate the problem in very explicit terms: using a specific set of low order invariant polynomials, generate a complete linearly independent set (over the real numbers) of invariant polynomials of degree less than or equal to the degree of the Sylvester polynomial. Determine whether the Sylvester polynomial lies in the span of this set. (Besides those invariants already mentioned, a reasonable set of invariants to pick for this application are those that can be generated by the symbolic method [Weyl] applied to the form det(H(t)).) By restricting attention to this specific problem we avoid the general problem that might be too hard to solve. This problem for the case n=3 is certainly not too hard -- we have already demonstrated a constructive algorithm that provides a polynomial of degree 36 (and degree 272 for n=4) in $\frac{1}{r}$ that is a sufficient condition that r be in S^{sing}. If possible, we would like to find a minimal set of lower order polynomials (e.g. two of order 18 for n=3) to do the same job.

It is worth noting that there are computer programs that generate and work with polynomial invariants using symbolic manipulation. The formulas 4.10a and 4.10b were computed by hand in about one-hour's time. That same computation could be performed by computer much faster (and with fewer errors at intermediate stages). A computer could also handle larger problems it could generate the Hessian invariants for the four, five, and six bock problems, for example. Moreover, there are deterministic algorithms for taking the (moderately large) set of invariants generated in this way and reduce it to a minimal set of generators. Finally, a computer could be used to determine whether a set of low-order invariants generate the Sylvester polynomial. Manual methods should be adequate to perform these tasks for the three and four block problems.

That summarizes where we are so far on invariants. In the near future we hope to determine at least in the case n=3 what the primitive factors of Sylvester's determinant polynomial are. Surely this problem has been worked on before [we have some leads, but no references yet]. The variety Z(det(H)) for n=3 is a singular cubic curve in P^2, and for n=4 it is a singular quartic surface in P^3. These two low-dimensional cases are reasonably well understood by algebraic geometers [Hartshorne]. It looks like algebraic geometry and the invariant polynomials could help us understand the low-dimensional problem and possibly lead us to a more efficient computational algorithm. To meet the improved efficiency challenge for n=4, it has to beat roughly 100 subroutine calls to the LINPACK routine dsvd with a 62×62 real matrix. An important practical goal is to find the real solutions efficiently by minimizing the computational overhead of filtering out the imaginary ones. For the few examples we have tested there seem to be relatively few real solutions, given the order of the underlying polynomials.

Last Minute Remarks: Given more time, we would have reworked the Canonical Form Theorem 3.1 so that the matrix U would be uniquely determined. For points in S^{sing} we now know how to do that, the approach is roughly as follows:

At $r \in S^{sing}$ the Singularity Condition implies that a U can be chosen so that

$$T_{n-1}\,[k^j\,(\alpha^j - \beta^j)] \tag{4.11}$$

(T_{n-1} from equation 3.10) is a real vector for all j. Furthermore, assuming the matrix rank is compatible, the complex coefficient matrix can be normalized so that the $n-1 \times n-1$ matrix:

$$T_{n-1}\,[k^2\,(\alpha^2 - \beta^2)\ \cdots\ k^n\,(\alpha^n - \beta^n)] \qquad (4.12$$

is the identity matrix. The phase of the vector k can be normalized so that the nonzero entry corresponding to the smallest index is positive. (Of course, there are exceptional cases where a different parametrization of a similar type must be used). Given these extra conditions, the canonical forms are reduced to their essential moduli.

Looking back at the 3-block problem, we see that this true canonical form leads to a manageable parametrization of the polynomial det(H(t)) in terms of α^j, β^j, andk. With all the free parameters absorbed into U, the polynomial expression can be parametrized directly by 11 real parameters (the c_{ijk} then become functions of these 11 geometric parameters). The dimension count breaks down as follows: only two real parameters are left in the vectors k^j $(\alpha^j - \beta^j)$; and the only other parameters needed are 9 more for the real part of the 2×2 matrix of cofactors in the lower right principal minor of the canonical form (the imaginary part of the cofactor matrix does not contribute to the determinant when the k^j $(\alpha^j - \beta^j)$ vectors are real). Of these 11, 3 disappear immediately because of the vanishing coefficients of the t_j^3 terms. Thus the seven algebraic coefficients c_{ijk} are conveniently parametrized by the geometric moduli fairly efficiently. The goal in this case would be to generate two more independent relations that must hold for $r \in S^{sing}$.

With the more efficient canonical form, perhaps the three-block problem can be solved by a direct computational approach.

Final Note: The author noticed the following just before the final deadline:

Observation 4.4: Suppose there is a point t^0 where det(H(t)) vanishes, nonvacuously, to first order. Expand the function det(H(t)) in a polynomial about t^0

$$\det(H(t)) = \sum_{j \geq i=1}^{n} q_{ij}(t)\,(t_i - t_i^0)(t_j - t_j^0)\,. \tag{4.13}$$

The polynomials $q_{ij}(t)$ of degree n-2 are not necessarily homogeneous. This globally valid expression for the polynomial in t has, in its expansion, expressions of varying total degree in t, yet it is known a priori to be homogeneous of degree n. Consequently, there is a set of n-1 nontrivial conditions R^k, one for each positive integer $k < n$, defined on the coefficients of the polynomials $q_{ij}(t)$, that must hold if the homogeneous polynomial of degree n-k in t is to vanish.

These conditions R^k are the key to THE EASIER RESULT OF EQUAL PRACTICAL VALUE - they should lead to a primary decomposition of the polynomial ideal generated by the gradient polynomials $\dfrac{\partial H(t))}{\partial t_j}$.

For k=1, the condition R^1 will be generated by a set of invariants R_i^1, each of which is the determinant of an $n \times n$ square matrix Λ_i^1 linear in the coefficients of the polynomials $q_{ij}(t)$. The condition R^1 is satisfied only if all these determinants vanish.

For higher k the picture is not so clear, but it looks as if R^k is also generated by a set of determinant functions: this time for matrices Λ_i^k of size equal to $s_n^k = \dfrac{n(n+1)\cdots(n+k-1)}{k!}$, the dimension of the space of homogeneous polynomials of degree k in n variables. Each matrix Λ^k will have entries that are polynomial functions (of degree $\leq k$?) in the coefficients of the polynomials $q_{ij}(t)$.

We believe can solve THE EASIER RESULT OF EQUAL PRACTICAL VALUE by the following algorithm:

1) For k=1,2 map the coefficients of the Hessian matrix polynomials $\frac{\partial^2 \det(H)}{\partial t_i \partial t_j}$ into the matrices Λ_i^k for the two values k=n-1 and n-2 (the Hessian polynomials are homogeneous of degree n-2, so lower order expressions in the expansion 4.13 vanish identically).

2) Evaluate the determinants of Λ_i^k to obtain the invariant functions $\phi_i^k(c)$, where c is the vector of coefficients c_J (J a multiindex of total degree n) of the original function det(H(t)).

3) Consider the set of invariant polynomials $\phi_i^k(c)$. It looks like the degrees are n (when k=1) and n(n+1) (when k=2), so the evaluation of these invariants is a reasonable numerical task (compared with n^n. The value r on which c depends is in S^{sing} (see Theorem 3.3) if and only if all the invariants $\phi_i^k(c)$ are zero.

We will investigate this idea further.

Notes for Section 4:

For background in invariant theory the reader is referred to [Weyl]. An older, more elementary book on invariants of (families of) quadratic forms is [Bromwich]. In some respects the approach we have taken here is an attempt to generalize the problem discussed in Bromwich's book but we are (so far)only working with families of rank-2 forms. An introduction to the modern theory of Algebraic Geometry is available in [Hartshorne]. A brief and interesting elementary chapter on real algebraic geometry can be found in [Milnor]. Finally, for anyone interested in working along these lines who is not familiar with the history of algebraic problems, we recommend [Dieudonne] as a general historical survey and the article by [Kleiman] as a brief survey of the related subject of Schubert's calculus. We did not have a chance to discuss here the Schubert-cycle interpretation of the Constructive Algorithm 3.1 -- that will have wait until later.

References for Section 4:

[Bromwich] T. J. Bromwich, Quadratic Forms and their Classification by Means of Invariant Factors, Hafner Publishing Co., Reprint from Cambridge University Press 1906.

[Dieudonne] Jean Dieudonne, History of Algebraic Geometry, Wadsworth 1985.

[Doyle] John Doyle, "Analysis of Feedback Systems with Structured Uncertainties," IEE Proceedings, vol. 129, Part D., No. 6, pp. 242-250, Nov. 1982.

[Hartshorne] Robin Hartshorne, Algebraic Geometry, Springer-Verlag, 1977.

[Kleiman] Steven L. Kleiman, "Rigorous Foundations of Schubert's Enumerative Calculus," from Mathematical Developments Arising from Hilbert Problems, F. Browder Ed., American Mathematical Society 1976.

[Milnor] John Milnor, Singular Points of Complex Hypersurfaces, Princeton University Press 1968.

[Packard 1] Andrew Packard, Ph. D. Thesis, University of California, Berkeley, Department of Mechanical Engineering 1988.

[Packard 2] Andrew Packard, Michael Fan, and John Doyle, "A Power Method for the Structured Singular Value," Preprint.

[Van der Waerden] Modern Algebra, Ungar Press, 1953.

[Weyl] Hermann Weyl, The Classical Groups, Princeton University Press, Second Edition, 1946.